

Recuperación de Información en Internet: Tema 1: Internet

P.O.P. Língua e usos profesionais

Miguel A. Alonso Jorge Graña Jesús Vilares

Departamento de Computación
Facultad de Informática
Universidade da Coruña

- 1 ¿Qué es Internet?
- 2 El flujo de información en Internet
- 3 Protocolos
- 4 Direcciones

¿Qué es Internet?

- Avance de la vigésima tercera edición del diccionario de la RAE:

internet. **1.** *amb. Inform.* Red informática mundial, descentralizada, formada por la conexión directa entre computadoras u ordenadores mediante un protocolo especial de comunicación.

ORTOGR. Escr. t. con may. inicial.

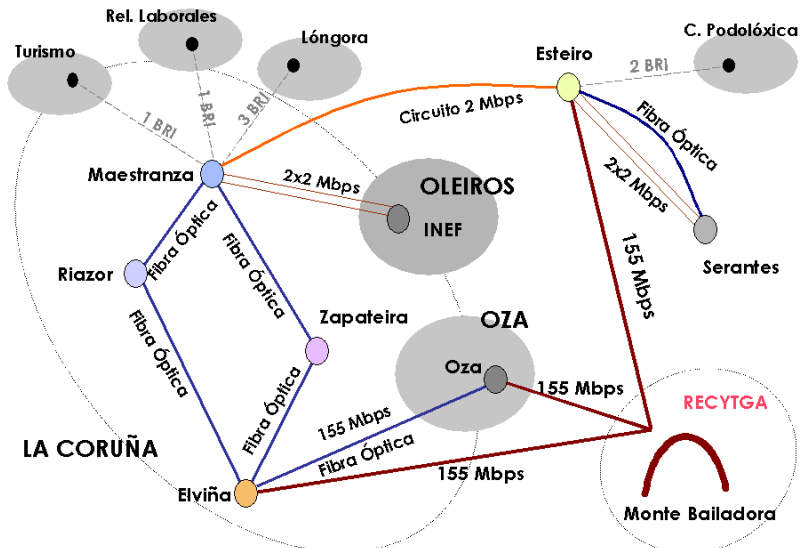
- Conceptos básicos:
 - es una “red de redes”, un mecanismo de **interconexión de redes** de computadoras
 - basada en un conjunto de **protocolos** denominado TCP/IP
 - garantiza que redes físicas heterogéneas funcionen como una red lógica única

Historia de Internet

- 1969 DARPA planifica la creación de una red que resista un ataque nuclear.
- 1972 Primera demostración pública de ARPANET, que funcionaba de forma distribuida sobre la red telefónica conmutada.
- 1983 ARPANET cambió el protocolo NCP por TCP/IP.
- 1986 NSF comenzó el desarrollo de NSFNET que se convirtió en la principal Red en árbol de Internet. Paralelamente, otras redes troncales en USA y Europa, públicas y comerciales, forman el esqueleto básico o "backbone" de Internet.
- 1989 Integración de los protocolos OSI en la arquitectura de Internet. En el CERN de Ginebra, un grupo de investigadores encabezado por Tim Berners-Lee, crearon el lenguaje HTML, basado en el SGML.
- 1990 El CERN implementa el primer cliente WWW y el primer servidor web.
- 2006 Internet alcanzó los mil cien millones de usuarios.

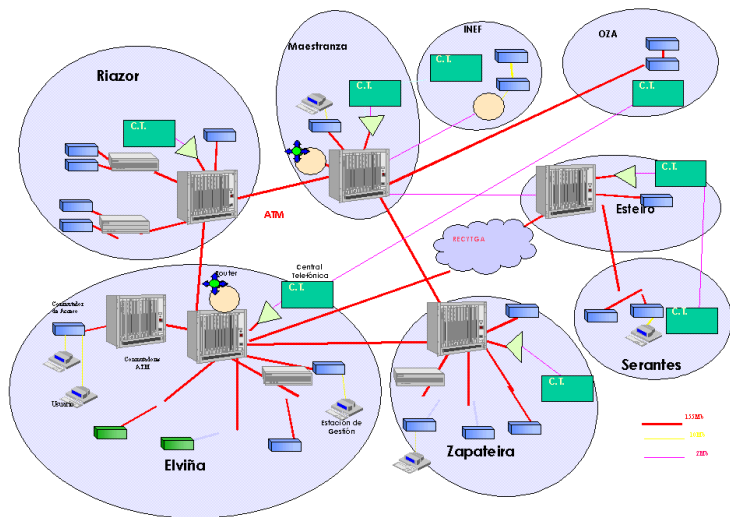
Red de redes

Internet conecta multitud de redes como las de la UDC:



Red de redes

Otra visión de la red de la UDC:



¿Qué se puede hacer en Internet?

Existen **protocolos** para cosas como:

- Realizar una **conexión remota** a otro ordenador (TELNET, SSH)
- **Transferir archivos** de un ordenador a otro (FTP)
- **Compartir ficheros** (NFS, SMB/CIFS)
- Gestionar el **correo electrónico** (SMTP, POP3, IMAP)
- Leer noticias (**news**) que escriben otros usuarios (NNTP)
- Leer y publicar **páginas web** (HTTP)
- **Charlar** con otros usuarios (IRC)
- **mensajería instantánea** (ICQ)
- ...

El flujo de información en Internet

● Arquitectura **cliente-servidor**

- Los servidores son ordenadores que ofrecen servicios.
- Los clientes son ordenadores que realizan peticiones al servidor para que éste realice algún servicio

● Paquetes de información

- La información se transmite en pequeños trozos llamados paquetes.
- Lo importante es la reconstrucción del mensaje emitido en el destino, no el camino seguido por cada paquete ni el orden de llegada.
- Si se destruye un nodo de la red, los paquetes encontrarán caminos alternativos.

TCP/IP

- Los **protocolos** son conjuntos de normas que regulan la transmisión de paquetes
- El protocolo **TCP** se encarga en el origen de fragmentar el mensaje emitido en paquetes. En el destino se encarga de reorganizar los paquetes para formar de nuevo el mensaje
- El protocolo **IP** direcciona los paquetes. Esto hace posible que los distintos paquetes que forman un mensaje pueden viajar por caminos diferentes hasta llegar al destino.
- Las redes de Internet usan TCP/IP para comunicarse entre sí

Pila de protocolos

A veces se usa el término TCP/IP para referirse a toda la “pila de protocolos” que hacen que las redes y los servicios que ofrecen funcionen

- Capa 1 – Nivel físico: Cable coaxial, fibra óptica, par trenzado, microondas, radio, ...
- Capa 2 – Nivel de enlace de datos: Ethernet, Gigabit Ethernet, Token Ring, FDDI, ATM, HDLC, WiFi, ...
- Capa 3 — Nivel de red: ARP, RARP, **IP (IPv4, IPv6)**, X.25, ICMP, IGMP, NetBEUI, IPX, Appletalk, ...
- Capa 4 — Nivel de transporte: **TCP**, UDP, SPX, ...
- Capa 5 — Nivel de sesión: NetBIOS, RPC, SSL, ...
- Capa 6 — Nivel de presentación: ASN.1, ...
- Capa 7 — Nivel de aplicación: SNMP, SMTP, NNTP, FTP, SSH, **HTTP**, SMB/CIFS, NFS, Telnet, IRC, ICQ, POP3, IMAP, ...

Las direcciones en Internet

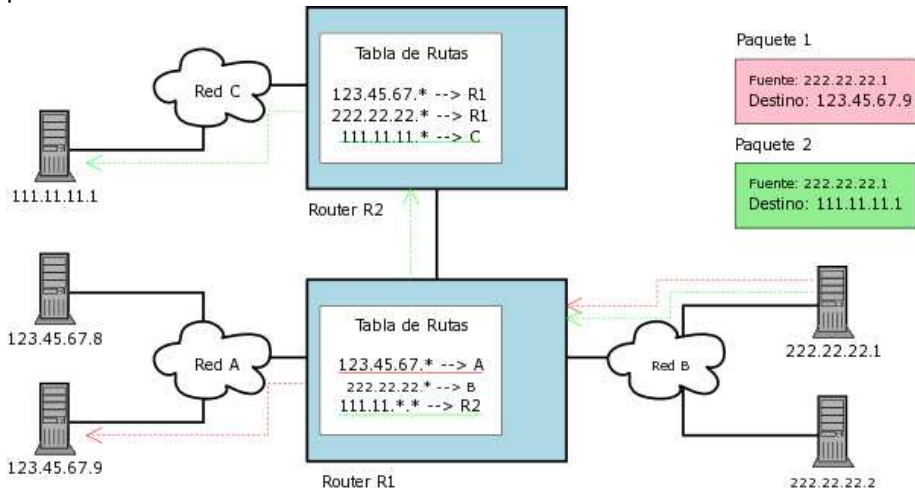
- Se emplean direcciones numéricas para identificar los ordenadores: las **direcciones IP**
 - Cuatro números, del 0 a 255, separados por puntos
 - Ejemplo: [193.144.51.17](#)
 - Pueden ser **estáticas** o **dinámicas** (DHCP)

Las direcciones en Internet

- Las direcciones IP se “traducen” a **nombres de dominio** y subdominio:
 - “palabras” separadas por puntos
 - La de más a la derecha indica el dominio de nivel superior
 - Hacia la izquierda de van indicando subdominios del dominio situado a su derecha
 - Normalmente la última “palabra” es el nombre de una máquina
 - Ejemplo: www.grupolys.org
 - Ejemplo: lys.dc.fi.udc.es
- Unos nombres pueden ser **alias** (sinónimos) de otros
- Puede haber muchos nombres que se refieran a la misma dirección IP
- La gestión de nombres la realiza el **DNS**, *Domain Name System*
- **ICANN**, *Internet Corporation for Assigned Names and Numbers*, es la organización que coordina la asignación de identificadores únicos en Internet: nombres de dominio, direcciones IP, números de puerto de los protocolos, ...

Enrutado

Los **routers**, enrutadores o encaminadores son los dispositivos de hardware para la interconexión de las redes de ordenadores:



Proxy

- La comunicación entre clientes y servidores no siempre es directa
- Frecuentemente aparece la figura del **proxy**, un ordenador que intercepta las conexiones de red que un cliente hace a un servidor de destino
- De forma genérica, podríamos definir un proxy como un ordenador en el que se delega la realización de ciertas conexiones
- El más común es el **proxy web**
- Pero puede haber proxies para muchos más protocolos, incluso para enrutar el tráfico de una red.

Cortafuegos

- Un cortafuegos o **firewall** es un elemento de hardware o software utilizado en una red de ordenadores para controlar las comunicaciones, permitiéndolas o prohibiéndolas según las políticas de red que se hayan definido.
 - Cortafuegos de **filtrado de paquetes**: filtran en función de las direcciones y puertos de origen y destino
 - Cortafuegos de la **capa de aplicación**: filtra en función del servicio (por ejemplo, denegando el acceso a ciertos servidores web)
 - Cortafuegos **personal**: software que filtra las comunicaciones entre un ordenador y la red
- Tratan de proteger una red de intrusiones y del acceso a información privada

Direcciones de usuarios y recursos

- Para identificar usuarios (por ejemplo en el **e-mail**) se usan direcciones con formato `usuario@nombre_de_máquina`
 - Ejemplo: alonso@udc.es
 - Ejemplo: pepito@lys.dc.fi.udc.es
- Cada recurso tiene una **URL** ó *Uniform Resource Locator*
 - El formato general es: `protocolo://máquina/directorio/fichero`
 - Ejemplo: <http://www.grupocole.org/cole/publications.html>
 - Pero pueden añadirse otros datos:
<http://pepiño:sfgu5hfq@www.misitio.com:80/cosas/fichero.html>